

文学研究論集

第50号 2019. 2

コーパスの英語教育への活用 ——難易度別英語教科書を例に——

The Application of Corpora to English Education ——A Study of Quantitative Analysis of Six Graded High School English Textbooks——

博士前期課程 英文学専攻 2017年度入学

櫻 井 大 暉
SAKURAI Daiki

【論文要旨】

本稿では、現在の日本の英語教育の傾向を取り上げ、その動向を把握した上で、コーパスを活用した日本の英語教育を概観し、コーパス言語学が英語教育に貢献するあり方を見ることで、櫻井（2018a, 2018b）における英語教科書を比較した分析結果が英語教育にどのように役立つのか、また今後の研究で有効なことについて考察している。

まず、現在の日本の英語教育に関して、今後の教育に向けて具体的な特徴として指摘されている大きな枠組みは「大学入試改革」「グローバル化」「テクノロジーの発達」の3点である。本論では、この3つの傾向に加えて、現在の英語教育において注目されている題材を取り上げて見ていく。

さらに、コーパスを活用した研究は、語法研究・社会言語学・辞書編纂・文体論・文学作品研究など多岐にわたるが、ここでは言語教育に焦点を当てて従来の活用方法を含め、近年注目をあびている CEFR の枠組みをコーパスに活用した研究についても見ていく。

最後に、櫻井（2018a, 2018b）での語彙及び文に関する分析結果が、近年の英語教育やコーパスを活用した英語教育の現状を鑑みて、どのように英語教育に資することができるかについて現時点での成果と今後の研究の展望も含め考察していく。

【キーワード】 言語教育、コーパスを活用した英語教育、難易度別英語教科書、語彙と文、CEFR

1. はじめに

本稿では、[1] 現在の日本の英語教育の動向を概観し、[2] コーパスを活用した日本の英語教育の貢献の事例を具体的に見て、[1]、[2] の状況を踏まえた上で、[3] 櫻井（2018a, 2018b）における難易度別英語教科書の分析が今後の英語教育に資する点を考察していく。

まず、[1] 現在の日本の英語教育の動向に関するキーワードとして「大学入試改革」「グローバル化」「テクノロジーの発達」の3点を取り上げる¹。以下本論では、これら3つの傾向とともに、新たな大学入試に向けてとり行われる取り組みについて見ていきたい。

また、[2] コーパスを活用した日本の英語教育について、従来より行われてきたものの一部として、石川（2008）をはじめ中村・堀田（2008）、中條・西垣・山保・天野（2011）、西垣・天野・吉森・中條（2011）、西垣・峰岸・中條（2012）、さらに大修館の『英語教育』で紹介されているものなどを具体的に見ていく。近年注目されている動向として、CEFR（Common European Framework of Reference for Languages）（ヨーロッパ言語共通参照枠）の枠組みを日本の英語教育環境に適用した CEFR-J を開発した東京外国語大学大学院の投野由紀夫氏のグループが行っている CEFR 関連の様々な研究にも焦点を当てて見ていきたい。

さらに、英語教育及びコーパスを活用した英語教育の現状を認識した上で、[3] 筆者が現在研究を進めている難易度別英語教科書の教材研究がこれからの英語教育にどのように貢献できるのかについて、語および語彙レベルの分析（語彙量、語彙多様性、語彙難易度、品詞構成比）、また文レベルの分析（文数、文長（各文中の語数）、構文）で得られた結果の一部の具体例を掘り下げて考察していきたい。

本論でいう英語教育への貢献とは、現場で教える教師や教材を開発する出版社など学習者に対して教育を提供する側が、教材を研究開発したり選択したりする際に役立つことであり、コーパス言語学の手法を使用した教材の評価方法を最も活用できるものの1つが、難易度別英語教科書の教材研究である。この研究が新たな教育の指針になり得ると考えている。

2. 現在の日本の英語教育

本節では、現在の日本の英語教育で注目されている事項を念頭に、大まかな枠組みと個別のテーマについて見ていく。

¹ 東京外国語大学大学院 根岸雅史氏（英語教育学）が『あしたの教育を考える教育最前線 これからの教育』（2018）の中で指摘した内容をまとめた。なお、上記3つのキーワードの2点目、3点目に関しては、文部科学省「今後の英語教育の改善・充実方策について 報告（概要）～グローバル化に対応した英語教育改革の五つの提言～」の中でも、「グローバル化の進展の中で、国際共通語である英語力の向上は日本の将来にとって極めて重要である。」「ICT 予算に係る地方財政措置を積極的に活用し、学校の英語授業における ICT 環境を整備。」のように取り上げられている。

2.1 英語教育に必要な基本的柱

冒頭にもあるが、英語教育の転換点の時期にこれからの英語教育を考えた上で重要な言葉に、「大学入試改革」「グローバル化」「テクノロジーの発達」の3つがある²。ここでは、この3つを取り上げて見ていく。

最初は、目前に迫った2020年大学入試改革である。今までとの大きな違いは、大学入試が4技能化し、これまで行われていた読解問題や文法・語彙問題中心でなくなる点だ。これまでの「読むこと」を問う問題だけではなく、「話すこと」「聞くこと」「書くこと」がこれから問題となる。そうすると、当然授業自体を今までとは違うものにしなければならず、そのためには教師や教材も入試に見合ったものにならなければならないため、教材に関してはコーパスを活用した教材研究の必要性が出てくるだろう。

次に、世界の急速なグローバル化だ。いよいよ2020年に開催される東京オリンピック・パラリンピックを契機に訪日外国人旅行者数が増え、一般の人でも英語の必要性が高くなり学校での英語教育もそれに対応できる力を養えるような授業が必要となってくる。また、グローバル化では、外国人と触れ合うことで異文化の良さや難しいところを実際に体験できるのも面白いところであり、教材などにおいても英語母語話者やそれ以外の英語話者のデータを集めたものと比較するなどして日本人と外国人の英語の使用差を検証し改善することもできるだろう。

最後の大きな枠組みは、近年のテクノロジーの急速な発達だ。とりわけ、自動翻訳や音声認識の技術の進歩が取り上げられる。教科書や新聞、論文などの論理構成が明白な文章には有効だが、人間同士の会話のような文脈による発話行為の翻訳は得意ではないようだ。よって、英語という言葉の特徴を把握し、実際に英語を話す人にしか分からないような機械では補えない人間特有の感覚を教師が授業で学習者に習得させることも重要である。

2.2 新大学入試

ここまで大きな枠組みについて見てきたが、今最も注目されているものの1つの新大学入試³に関する取り組みを見ていきたい。

大学入試改革では、「学力の3要素」⁴である「(1)知識・技能、(2)思考力・判断力・表現力、(3)主体性を持って多様な人々と協働して学ぶ態度」について、多面的・総合的評価する入試になろうとしている。2020年度に「大学入学共通テスト」が開始、2024年度には新学習指導要領を前提に更に改革する計画となっている。英語においては、2技能評価（「読む」「聞く」）から4技能評価になる。

² 注1同様。

³ 東京外国語大学大学院 根岸雅史氏（英語教育学）が『三省堂高校英語教育』（2018）で指摘した内容をまとめた。

⁴ 文部科学省の学習指導要領「生きる力」第1章 言語活動の充実に関する基本的な考え方に示されている。

大学入試センター試験の代わりの大学入学共通テストでは、知識重視の傾向から思考力・判断力・表現力を測ることとなっている⁵。英語においては、外部試験である認定試験を利用し、「スピーキング」や「ライティング」が測定される。

また、認定試験⁶の利用に関して、「日本国内において、…広く実施されている実績がある」「英語4技能の全てを極端な偏りなく評価する」「高等学校学習指導要領との整合性が図られている」「CEFR との対応関係並びにその根拠となる検証方法及び研究成果等が公表」されている、などの基準を満たすと判断されたものが認定試験として認められる。認定されたテストは、IETLS, GTEC, TEAP, TEAP CBT, TOEFL iBT, TOEIC, ケンブリッジ英語検定、実用英語技能検定（公開会場実施の4技能同時受験型）である。これらのテストはそれぞれの評価の基準が異なるため、CEFR の枠組みを用いて比較した一覧表が文部科学省より公開されている。（図1）

各資格・検定試験とCEFRとの対照表

文部科学省（平成30年3月）

CEFR	ケンブリッジ 英語検定	実用英語技能検定 1級-3級	GTEC Advanced Basic Core CBT	IELTS	TEAP	TEAP CBT	TOEFL iBT	TOEIC L&R/ TOEIC S&W
C2	230 200 (230) (210)			9.0 8.5				
C1	199 180 (190)	3299 2600 (3299) (2630)	1400 1350 (1400)	8.0 7.0	400 375	800	120 95	1990 1845
B2	179 160 (170)	2599 2300 (2599) (2304)	1349 1190 (1280)	6.5 5.5	374 309	795 600	94 72	1840 1560
B1	159 140 (150)	2299 1950 (2299) (1980)	1189 960 (1080)	5.0 4.0	308 225	595 420	71 42	1555 1150
A2	139 120 (130)	1949 1700 (1949) (1728)	959 690 (840)		224 135	415 235		1145 625
A1	119 100 (100)	1699 1400 (1699) (1456)	689 270 (270)					620 320

○ 表中の数値は各資格・検定試験の定める試験結果のスコアを指す。スコアの記載がない欄は、各資格・検定試験において当該欄に対応する能力を有していると認定できないことを意味する。

【図1 各資格・検定試験と CEFR との対照表（文部科学省，平成30年）】

英語教師にとっては当然経験したことのない変革なので、これからこの試験形式を経験する学習者のための授業についても、教師が説明するだけの従来型の授業ではなく、学習者が主体的に授業の中で英語を使用する機会を作るような授業作りをする必要があるだろう。

⁵ 文部科学省 高大接続改革「大学入学共通テスト」についての中で指摘されている。
(http://www.mext.go.jp/a_menu/koutou/koudai/detail/1397733.htm)

⁶ 文部科学省 大学入試英語成績提供システム参加要件を満たした試験のことを指す。

3. コーパスを活用した日本の英語教育

ここでは、コーパスを用いた様々な活用法及び近年急速に研究が進んでいる CEFR-J について見ていく。

3.1 様々な活用法

まず、英語教育にコーパスを活用した研究にはどのようなものがあるのか見ていく。まず、櫻井 (2018a) の主要な先行研究にもなった石川 (2008) では、日本および韓国の英語教科書、その参照コーパスとなる1990年代の英語を代表する FROWN, FLOB コーパス (各100万語)、英米映画のセリフデータ (計130万語) を利用し、「語彙量・多様性・難度・品詞構成比・意味領域」等に関して分析が行われている。結論として、日本の教科書の語彙量の少なさは圧倒的で、韓国の教科書の延べ語数の50%, レマ数の44.7%にとどまり、標準英語のレマ総数に占めるカバー率は13.2%にとどまる。さらに、日本の教科書の品詞分布は英語の標準からかなり逸脱している。例えば、I の過剰使用、動詞不定詞形の過剰使用を抑制し、動詞時制の多様性を確保し、接続詞を増やす必要性がある。また、日本の教科書の意味領域の分布は英語の標準から逸脱している。日本の教科書は日常的・生活的意味領域に偏っており、社会・経済・政治などの領域が弱い。これらの結果が出ている。

また、土村 (2014) では、1963年から2014年に至る英語教科書 2 種類 (計36冊) をデータとし、語法やコロケーション、タグ付けされた品詞や意味領域の頻度の観点から、教科書と英語母語話者の使用する英語の相違を年代別に分析している。さらに、中村・堀田 (2008) では、中学校の教科書 3 学年分 7 種類 (計21冊) をデータとし、平均文長や TTR の推移、語彙の伸長度、使用頻度等に関して研究が行われている。

以下では、コーパスを活用した英語教育では主要なテーマの 1 つであるデータ駆動型学習 (DDL: Data-driven Learning) について見ていく。DDL とは、学習者自身がコーパスで単語・語句や文法項目の規則性を見出し、コンコダンスラインなどを観察して、実際に調査したい語彙や文法の使われ方を帰納的に発見し学習する方法である。

DDL の実際の活用例を見ると、中條・西垣・山保・天野 (2011) では、DDL と呼ばれる学習者自らがコーパスを検索・観察して利用する方法について、日本・中国・韓国・台湾の学校教科書を対象に難易度を測定し、初級レベル学習者でも使用できるようなコーパス作成を目指す研究が行われている。西垣・天野・吉森・中條 (2011) では、DDL を中・高生の学習者の指導に活用するために、日本・中国・韓国・台湾の英語教科書コーパスを利用して、学習者の英語力レベルにあったコーパスを用い、さらに中・高生のための DDL ワークシートを開発して、中・高の通常の授業での DDL の導入を試みている。西垣・峰岸・中條 (2012) では、語彙・文法の新しい指導法として、生徒に文法・語法の規則を帰納的に発見させるワークシートを作成し、四カ国英語教科書コーパスを利用して、中・高生にわかりやすいコンコダンスラインの提示方法の指導について研究されている。

3.2 CEFR-Jの枠組み

まず、CEFR について見ていく。CEFR とは、「語学シラバスやカリキュラムの手引きの作成、学習指導教材の編集、外国語運用能力の評価のために、透明性が高く、分かりやすい、包括的な基盤を提供するものとして、20年以上にわたる研究を経て、2001年に欧州評議会が発表した、外国語の学習者の習得状況を示す際に用いられる枠組み」⁷である。図2が示すように6つのレベルに分けてそれぞれのレベルでできることを規定しているものである。

次に、東京外国語大学大学院の投野由紀夫氏のグループが研究を行っている CEFR-J について見ていく。CEFR-J とは、ヨーロッパ言語共通参照枠（CEFR）に準拠した汎用枠のことを指す。日本の英語教育が初級レベルで苦闘している現状に鑑み、CEFR の6レベルを Pre-A1～C2 までにすることで12レベルに細分化、基礎レベルを手厚くした独自の CAN-DO ディスクリプタを有している。2018年現在、CEFR-J 科研⁸は4期目を迎え、第3期で整備した CEFR-J レベル別言語材料をもとにタスクとテスト開発の研究に移行している。CEFR-J の CAN-DO リストと語彙・文法・テキストのレベル別資料が整うことで、小学校（Pre-A1～A1.2相当）、中学校（A1.3～A2.2相当）で具体的なシラバス設計を試みることができる。CEFR-J Grammar Profile については成城大学の石井康毅氏が、CEFR-J Text Profile については九州大学の内田論氏が分担して研究を担っている⁹。

まず、CEFR-J Grammar Profile とは、英語の文法項目の各種 CEFR(-J) レベル別コーパスにおける頻度データをもとに、各 CEFR-J レベルに対応する文法項目の導入時期を特定するための取り組みのことである。日本人英語学習者に各項目をどの学習・習熟段階で指導すべきかの判断に資することを目的とし、CEFR-J 科研チームによって開発されている。

また、CEFR-J Text Profile とは、CEFR レベル別の海外の EFL コースブックのコーパスをもとに、各 CEFR レベルにおける語彙および構文の特徴を明らかにすることを目的に作成されたものである。A1, A2, B1, B2, C の5段階で、A1 レベルの語がどの程度の割合で出現するか、文の平均的な長さはどの程度か、1文あたりに名詞句はどの程度出現するか、などについての基本統計量（平均・分散など）を示している。

本著者の研究は、学習者のレベル別に難易度が異なる構成の教科書を取り扱っている点で、CEFR(-J) の目的・分析の考え方に共通する部分がある。今後の研究の参考にしたい。

⁷ 文部科学省「大学入試英語成績提供システムへの参加要件を満たしている資格・検定試験と CEFR との対照表について」を参照。

⁸ CEFR-J は科学研究費基盤研究(A)を3回計12年間継続して構築した本格的な CEFR 準拠の枠組である。CEFR を細分化した12段階のレベルそれぞれに CAN-DO リストを専門家の研修を受けたチームで執筆、CAN-DO の妥当性検証を科学的に行っている。(CEFR-J 公式ホームページより。)

⁹ 石井康毅 (2016). 「英文中の文法項目頻度調査のための項目選定と英文からの抽出法—CEFR-J の枠組みでの研究—」『社会イノベーション研究』, 11, 2016, 107-122. や内田論 (2017). 「English Vocabulary Profile を語彙指導に活用する」『英語教育』, 2, 2017, 32-34. などを参照。

熟練した言語使用者	C2	聞いたり読んだりした、ほぼ全てのものを容易に理解することができる。いろいろな話し言葉や書き言葉から得た情報をまとめ、根拠も論点も一貫した方法で再構築できる。自然に、流暢かつ正確に自己表現ができる。
	C1	いろいろな種類の高度な内容のかなり長い文章を理解して、含意を把握できる。言葉を探しているという印象を与えずに、流暢に、また自然に自己表現ができる。社会生活を営むため、また学問上や職業上の目的で、言葉を柔軟かつ効果的に用いることができる。複雑な話題について明確で、しっかりとした構成の、詳細な文章を作ることができる。
自立した言語使用者	B2	自分の専門分野の技術的な議論も含めて、抽象的な話題でも具体的な話題でも、複雑な文章の主要な内容を理解できる。母語話者とはお互いに緊張しないで普通にやり取りができるくらい流暢かつ自然である。幅広い話題について明確で詳細な文章を作ることができる。
	B1	仕事、学校、娯楽などで普段出会うような身近な話題について、標準的な話し方であれば、主要な点を理解できる。その言葉が話されている地域にいたときに起こりそうな、たいていの事態に対処することができる。身近な話題や個人的に関心のある話題について、筋の通った簡単な文章を作ることができる。
基礎段階の言語使用者	A2	ごく基本的な個人情報や家族情報、買い物、地元の地理、仕事など、直接的関係がある領域に関しては、文やよく使われる表現が理解できる。簡単で日常的な範囲なら、身近で日常の事柄について、単純で直接的な情報交換に応じることができる。
	A1	具体的な欲求を満足させるための、よく使われる日常的表現と基本的な言い回しは理解し、用いることができる。自分や他人を紹介することができ、住んでいるところや、誰と知り合いであるか、持ち物などの個人的情報について、質問をしたり、答えたりすることができる。もし、相手がゆっくり、はっきりと話して、助けが得られるならば、簡単なやり取りをすることができる。

(出典) プリティッシュ・カウンシル、ケンブリッジ大学英語検定機構

【図2 CEFR が示す 6 段階の共通参照レベル】

4. 難易度別英語教科書の教材研究への応用

最後に、英語教育やコーパスを活用した英語教育の現状を念頭に置いた上で、筆者の現在の研究テーマである「難易度別英語教科書の教材研究」がこれからの英語教育に貢献し得る点について、櫻井（2018a, 2018b）の分析結果の一部を取り上げて考察していく。また、表1には櫻井（2018a, 2018b）で分析対象とした英語教科書を示す。

櫻井（2018a, 2018b）における研究でも触れたように、英語教科書は、英語学習者にとって最も重要な学習素材（教材）の1つであり、出版社によっては各学校（採用校）のレベルに合わせて難易度別に教科書が作成されている。すなわち、三省堂、東京書籍、数研出版、第一学習社の4社¹⁰では、上位校、中堅校、下位校に合わせて3つのレベルの教科書が用意されている。表1中

【表1. 櫻井（2018a, 2018b）において分析に用いた難易度別英語教科書一覧】

出版社	難易度	教科書〈コミュニケーション英語〉		
		高校1年	高校2年	高校3年
東京書籍	レベル1	『All Aboard! I 』(A1) ¹¹	『All Aboard! II 』(A2)	『All Aboard! III 』(A3)
	レベル2	『Power On I 』(PO1)	『Power On II 』(PO2)	『Power On III 』(PO3)
	レベル3	『PROMINENCE I 』(PR1)	『PROMINENCE II 』(PR2)	『PROMINENCE III 』(PR3)
三省堂	レベル1	『VISTA I 』(V1)	『VISTA II 』 ¹² (V2)	
	レベル2	『MY WAY I 』(M1)	『MY WAY II 』(M2)	『MY WAY III 』(M3)
	レベル3	『CROWN I 』(C1)	『CROWN II 』(C2)	『CROWN III 』(C3)

¹⁰ 著者の調査では、現在この4社が3つのレベルの教科書を用意している。

¹¹ 表中の（ ）内は、以降の文・表中で使用される教科書の略号となっている。

¹² VistaはVista I、IIの2冊で3年間分の学習範囲設定となっている。

の難易度は、レベル 1（初級）・レベル 2（中級）・レベル 3（上級）の順に上がっていく設定となっている。今回は、その中から 2 つの出版社（東京書籍、三省堂）の全 6 種のコミュニケーション英語を分析対象とした。

4.1 語彙量・語彙多様性

まず、櫻井（2018a）において最も特徴的な結果となった「語彙量・語彙多様性」を取り上げて考察していく。各教科書をレベル別（レベル 1～3）に分けたものを表 2 に示す。ここでは、各数値を詳細に見ていくことはせずにレベルごとの傾向を捉えて考察していく。

最初に、分析した項目の概要について簡単に触れておく。各種テキストやコーパスの総語数を数える時に、延べ語数と異なり語数という言葉を使用する。延べ語数とは、単語をふつうに数えた時の語数（テキストに現れる語の累計数、コーパス総語数）のことで、異なり語数とは、単語の種類として数えた時の語数のことである。例えば、*Kanna went to her school to talk to her boyfriend.* という文を見てみると、延べ語数は「10語」となるが、異なり語数は *to* と *her* が重複しているため「7語」という計算になる。また、TTR (Type/Token Ratio) とは、異なり語数を延べ語数で割った百分率のことで、テキストの中にどのくらい異なる語が含まれているか（テキストの語彙多様性）を表すものとなっている。Guiraud 値とは、異なり語数を延べ語数の平方根で割ったもので、延べ語数の増加によって値が過度に低下しないように意図されたものとなっている。実際に語彙多様性を比較する時には Guiraud 値の数値を参照する。レマ (lemma) とは、活用形、綴り字の違いを問わず、語幹と語類を同じくする各種の表記形を含む基準形のことである。例えば *play* という単語を見てみると、表記形は *play*, *plays*, *played*, *playing* があるが、レマは *PLAY* という 1 つの種類に分類される。

それでは、どのような点が英語教育に貢献し得るのか考察していく。英語教育への貢献というと、学習者にとっての利益をはじめ様々な意味合いを持つが、本論でいう英語教育への貢献とは、現場で教える教師や教材を開発する出版社など学習者に対して教育を提供する側が、教材を開発したり、選択したりする際に役立つことである。

まず、延べ語数に関して、表 2 を見るとレベルが上がるごとに語数が多くなることが分かる。また、異なり語数やレマ語数が判明することで、延べ語数では難易度別に語数が多くなっているが、異なり語数やレマ語数では延べ語数ほど上昇の割合は高くないということが分かる。TTR や Guiraud 値を算出することで、語彙量だけでは分からない語彙多様性が判明し、難易度別に確かに語彙量は上がっているが語彙の種類はどうなのだろうかということを知りたい時に大いに役立つものとなっている。ここでは、難易度別に語彙多様性が上昇する傾向となっている。コーパスの分析により学習者の習熟度に合わせた総語数や異なり語数、レマ語数、Guiraud 値を導き出すことによって、教科書を含めた教材や試験全般に関する学習者のレベルに見合った教材を作成する一助となるという意味で、難易度別英語教科書の教材研究は英語教育に資することができるのではないかと

【表 2. 櫻井（2018a）における難易度別英語教科書の語彙量・語彙多様性〈レベル別〉】

【レベル 1】	【三省堂】		【東京書籍】		
	V1	V2	A1	A2	A3
異なり語数（レマ）	596	590	520	727	719
異なり語数（ワード）	691	665	585	827	819
延べ語数	1809	1914	1515	2630	2596
TTR	38.2%	34.7%	38.6%	31.4%	31.5%
Guiraud 値	16.25	15.20	15.03	16.13	16.07
TTR（レマ）	32.9%	30.8%	34.3%	27.6%	27.7%

【レベル 2】	【三省堂】			【東京書籍】		
	M1	M2	M3	PO1	PO2	PO3
異なり語数（レマ）	857	1013	1262	865	1111	994
異なり語数（ワード）	1014	1181	1524	1021	1335	1175
延べ語数	3772	4375	6261	3627	4897	4314
TTR	26.9%	27.0%	24.3%	28.1%	27.3%	27.2%
Guiraud 値	16.51	17.86	19.26	16.95	19.08	17.89
TTR（レマ）	22.7%	23.2%	20.2%	23.8%	22.7%	23.0%

【レベル 3】	【三省堂】			【東京書籍】		
	C1	C2	C3	PR1	PR2	PR3
異なり語数（レマ）	1209	1408	1862	1325	1418	2567
異なり語数（ワード）	1520	1718	2362	1606	1767	3252
延べ語数	7053	7306	12078	6406	7280	15494
TTR	21.6%	23.5%	19.6%	25.1%	24.3%	21.0%
Guiraud 値	18.10	20.10	21.49	20.07	20.71	26.13
TTR（レマ）	17.1%	19.3%	15.4%	20.7%	19.5%	16.6%

考えている。

具体的により良い教材作成に役立てることを考えた時、語彙量ではレベルが上がるにつれて延べ語数は格段に増加しているが、異なり語数は、延べ語数が増加する割合ほどは増加していないことが分かる。全体の延べ語数としての語彙量は増加しても単語の種類別に見た異なり語数としての語彙量は比例して増加していないことになるので、延べ語数を増やすのであれば異なり語数も相応に増やした方が語彙の多様性をより確保できるような教材を作成することができるのではないだろうか。一方、テキスト内に同じ単語が多く出現した方が様々な環境でその単語の使われ方を習得し、定着を図ることができるという点で利点があるということもできるので、バランスを考慮した構成にする必要がある。

4.2 文数・文長

ここでも、櫻井（2018b）において最も特徴的な結果となった「文数・文長」を取り上げて考察していく。各教科書をレベル別（レベル1～3）に分けたものを表3に示す。同様に、各数値を詳細に見ていくことはせずにレベルごとの傾向を捉えて考察していく。

本稿では、「文」とは、ピリオド・クエスションマーク・エクスクラメーションマーク（., ?, !）のマーカで区切られた文字列とする。また、文長について、一文ずつに分けたものの語数の平均を各教科書ごとに算出している。本稿では、「文長」は各文（1文）内の単語の数と定義する。「語」とは総語数〔延べ語数〕で数えられるものと定義し、「I' m'」などの短縮形（縮約形）も1語としている。

同様に、どのような点が英語教育に貢献し得るのか考察していく。一般的には語彙レベルだけに注目しがちかもしれないが、文の数を分析し平均文長が分かることでそのテキストが全体としてどのぐらいの単語の長さで1文が構成されているか分かる。文の複雑さに関しては一概には言えないが、普通1文の語彙数が多ければ自ずと文の難しさも増す可能性も高いだろう。¹³ 難易度別英語教科書においては、レベルが上がるにつれて平均文長が長くなることが分かっている。ここでも、コーパスの分析を利用することにより、学習者の習熟度に合わせた教材を作成する指標を提供できる点で英語教育に貢献できるのではないかと考えている。

こちらにも具体的により良い教材作成に役立てることを考えた時、各教科書の総文数と文長が分か

【表3. 櫻井（2018b）における難易度別英語教科書の総文数・文長〈レベル別〉】

【レベル1】	【三省堂】		【東京書籍】		
	V1	V2	A1	A2	A3
総文数	217	212	163	249	209
文長（各文中の語数）	8.16	8.91	9.25	11	12.4

【レベル2】	【三省堂】			【東京書籍】		
	M1	M2	M3	PO1	PO2	PO3
総文数	385	400	446	328	400	303
文長（各文中の語数）	9.84	11	14.1	11	12.2	14.17

【レベル3】	【三省堂】			【東京書籍】		
	C1	C2	C3	PR1	PR2	PR3
総文数	564	517	770	541	430	855
文長（各文中の語数）	12.4	14.1	15.6	12	16.8	17.86

¹³ 「文の複雑さ」については、文の長さ（文中の語彙数）や使用されている文法項目の種類（例えば、関係詞や接続詞などの使用により文構造が分かりにくくなる）などが関係すると考えられる。文長も含め、文法項目などの「文の複雑さ」に影響する要因を今後の分析で具体的に明らかにしていきたい。

ると、総文数と文長両方のバランスを考慮してレベル別により適した教材を作成することができ
る。例えば、ある教科書は総文数は多いのに文長は短いなら比較的短い文（単語が少ない）が多く
存在する教材ということになり、逆の場合、比較的長い文（単語が多い）が多く存在する教材とい
うことになる。その出版社において教科書全体のボリュームが決まっている場合は特に、文の複雑
さの1つの要因と考えられる文長をコントロールすることで、学習者のレベルに合わせたより良
い教材を作成することに繋がるのではないか。

本稿では文の複雑さの指標として従来から使用されている平均文長を分析してきたが、同様の指
標としていわゆる「リーダビリティ」スコア (RS) がある。RS には複数の提案がなされているが、
ここではそのうち ARI (Automated Readability Index)¹⁴ および CLI (Coleman-Liau Index)¹⁵ を使
用し、教科書データの1部を分析した結果を示す。一般的に1つの教科書において最初のレッス
ンから最後のレッスンになるにつれて難易度は上がると考えられる。そこで、試験的に *PROMI-
NENCE I* (PR1) の最初 (Lesson 1) と最後 (Lesson 10) の本文を Word Level Checker¹⁶ を使
用し分析した。Lesson 1 では、ARI: 4.8, CLI: 8.0 となり、Lesson 10 では、ARI: 8.6, CLI: 10.5 と
なる。¹⁷ Lesson 1 内での差は生じているものの、Lesson 1 と比較して Lesson 10 の値は両指標とも
高くなることが分かる。この2つの RS には平均文長、平均語長の2種類の変数しか使用されてい
ないが、その他の要因となり得る文法項目などの指標も考慮した分析もしていきたいと考えている。

5. 終わりに

以上、本論では、英語教育、コーパスを活用した英語教育を概観し現状を把握した上で、難易度
別英語教科書の教材研究への応用について取り上げコーパスが英語教育に貢献できる点について考
察してきた。先にも述べたように、難易度別英語教科書の教材研究の利用により、英語教育を提供

¹⁴ ARI の算出式は、以下の通りである。

$$\text{ARI} = (4.71 \times \text{CPW}) + (0.5 \times \text{WPS}) - 21.43$$

where,

CPW = average number of characters per word (= total number of C divided by that of W)

WPS = average number of words per sentence (= Total number of W divided by that of S)

¹⁵ CLI の算出式は、以下の通りである。

$$\text{CLI} = (5.89 \times \text{CPW}) - (0.3 \times (100 \div \text{WPS})) - 15.8$$

where,

CPW = average number of characters per word (= total number of C divided by that of W)

WPS = average number of words per sentence (= Total number of W divided by that of S)

詳細は、染谷泰正 (2009). 『オンライン版「英文語彙難易度解析プログラム」(Word Level Checker) の概
要とその応用可能性について』に記載されている。

¹⁶ Word Level Checker は染谷泰正氏が青山学院大学在籍時に開発。本稿執筆時点で、染谷氏は関西大学在籍
となっているが、Word Level Checker は青山学院大学によってサービスが継続されている。

¹⁷ Lesson 1 内では ARI の値と CLI の値の差が大きい、これは Lesson 1 の語彙量が少ないために出た結果
と考えられる。

する側が学習者にとってより良い教材や教育指標となるような分析結果を活かし、英語教育へ繋げていきたい。

今後特に、難易度別英語教科書を提供している残りの出版社（数研出版，第一学習社）のデータを加え，CEFR(-J) の取り組みも大いに参考にし，難易度別英語教科書の教材研究を通じてコーパスが英語教育に貢献できる様々な事項について研究を続けていきたい。

【参考文献】

Aoyama Gakuin University Word Level Checker.

(http://someya-net.com/wlc/index_J.html)

中條清美・西垣知佳子・山保太力・天野孝太郎 (2011). 「英語初級者向けコーパスデータとしての教科書テキストの適性に関する研究」, 『日本大学生産工学部研究報告 B』, 44, 2011, 13-23.

CEFR-J. 『What is CEFR-J?』 (<http://www.cefr-j.org/cefrj.html>)

Council of Europe (2011). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

石井康毅 (2016). 「英文中の文法項目頻度調査のための項目選定と英文からの抽出法—CEFR-J の枠組みでの研究—」 『社会イノベーション研究』, 11, 2016, 107-122.

石井康毅 (2018). 「CEFR-J Grammar Profile とは何か」 『英語教育』, 6, 2018, 東京：大修館書店, 35-36.

石川慎一郎 (2008). 「コーパスと教材研究」 『英語コーパスと言語教育—データとしてのテキスト—』: 159-183. 東京：大修館書店.

文部科学省 (2009). 『高等学校学習指導要領「生きる力」』.

文部科学省 (2014). 『今後の英語教育の改善・充実方策について 報告（概要）～グローバル化に対応した英語教育改革の五つの提言～』.

(http://www.mext.go.jp/b_menu/shingi/chousa/shotou/102/houkoku/attach/1352463.htm)

文部科学省 (2017). 『高大接続改革「大学入学共通テスト」について』.

(http://www.mext.go.jp/a_menu/koutou/koudai/detail/1397733.htm)

文部科学省 (2017). 『大学入試英語成績提供システム参加要件（大学入試センター策定）の公表について』

(http://www.mext.go.jp/b_menu/houdou/29/11/1398169.htm)

文部科学省 (2018). 『大学入試英語成績提供システムへの参加要件を満たしている資格・検定試験と CEFR との対照表について』.

(http://www.mext.go.jp/b_menu/houdou/30/03/_icsFiles/afieldfile/2018/03/26/1402610.pdf)

中村純作 (2008). 「教科書コーパスから何が見えるか—方法論と中学校英語教科書の場合—」 中村純作・堀田秀吾 『コーパスと英語教育の接点』: 121-150. 東京：松柏社.

根岸雅史 (2018). 「これからの英語教育—三つの波」 『あしたの教育を考える教育最前線 これからの教育』: 8-9. 東京：三省堂.

(https://tb.sanseido-publ.co.jp/wp-sanseido/wp-content/themes/sanseido/pdf/vanguard_1.pdf)

根岸雅史 (2018). 「新「大学入試」とこれからの授業」 『三省堂高校英語教育』: 2-5. 東京：三省堂.

(https://tb.sanseido-publ.co.jp/wp-sanseido/wp-content/themes/sanseido/pdf/h-school/hs-english/2018_Magazine_kokoeigo.pdf)

西垣知佳子・天野孝太郎・吉森智大・中條清美 (2011). 「中・高生のためのコンコーダンス・ラインを利用したデータ駆動型英語学習教材の開発の試み」, 『千葉大学教育学部研究紀要』, 59, 2011, 235-240.

西垣知佳子・峰岸識子・中條清美 (2012). 「中学・高校の英語教育におけるデータ駆動型学習に基づく帰納的学習の実践的研究」, 『千葉大学教育学部研究紀要』, 60, 2012, 419-426.

櫻井大暉 (2018a). 「難易度別英語教科書 6 種の計量的教材研究—【1】語彙の考察—」, 『文学研究論集』, 第 48 号, 19-36.

- 櫻井大暉 (2018b). 「難易度別英語教科書 6 種の計量的教材研究—【2】文の考察—」, 『文学研究論集』, 第49号, 1-17.
- 染谷泰正 (2009). 『オンライン版「英文語彙難易度解析プログラム」(Word Level Checker) の概要とその応用可能性について』
- 投野由紀夫 (2018). 「CEFR-J 活用研究の最新動向」『英語教育』, 6, 2018, 東京: 大修館書店, 34.
- 土村成美 (2014). *Quantitative Analysis of High School English Textbooks —Diachronic Change and Difference from Native Speakers—*. 修士論文. 明治大学.
- 内田 諭 (2017). 「English Vocabulary Profile を語彙指導に活用する」『英語教育』, 2, 2017, 32-34.
- 内田 諭 (2018). 「CEFR-J Text Profile とは何か」『英語教育』, 6, 2018, 東京: 大修館書店, 36-37.

【付録】 コーパスに用いた教科書リスト

題 名	対象学年	編著者	出版年	出版社	ページ数
VISTA English Communication I	高校 1 年	金子朝子他	2013	三省堂	136
VISTA English Communication II	高校 2・3 年	金子朝子他	2014	三省堂	128
MY WAY English Communication I	高校 1 年	森住衛他	2013	三省堂	168
MY WAY English Communication II	高校 2 年	森住衛他	2014	三省堂	176
MY WAY English Communication III	高校 3 年	森住衛他	2015	三省堂	152
CROWN English Communication I	高校 1 年	霜崎實他	2013	三省堂	192
CROWN English Communication II	高校 2 年	霜崎實他	2014	三省堂	216
CROWN English Communication III	高校 3 年	霜崎實他	2015	三省堂	192
All Aboard! English Communication I	高校 1 年	清田洋一他	2013	東京書籍	128
All Aboard! English Communication II	高校 2 年	清田洋一他	2014	東京書籍	152
All Aboard! English Communication III	高校 3 年	清田洋一他	2015	東京書籍	103
Power On English Communication I	高校 1 年	浅見道明他	2013	東京書籍	168
Power On English Communication II	高校 2 年	浅見道明他	2014	東京書籍	175
Power On English Communication III	高校 3 年	浅見道明他	2015	東京書籍	127
PROMINENCE English Communication I	高校 1 年	田辺正美他	2013	東京書籍	167
PROMINENCE English Communication II	高校 2 年	田辺正美他	2014	東京書籍	183
PROMINENCE English Communication III	高校 3 年	田辺正美他	2015	東京書籍	191